DOCUMENT RESUME

ED 458 464 CG 031 325

AUTHOR Campbell, David
TITLE The History and Development of the Campbell Interest and Skill Survey (CISS).
PUB DATE 2001-08-27
NOTE 24p.; Paper presented at the Annual Conference of the American Psychological Association (109th, San Francisco, CA, August 24-28, 2001).
PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Attitude Measures; *Career Counseling; Career Guidance; *History; *Interest Inventories; *Measures (Individuals); Skill Analysis
IDENTIFIERS Campbell Interest and Skill Survey

ABSTRACT

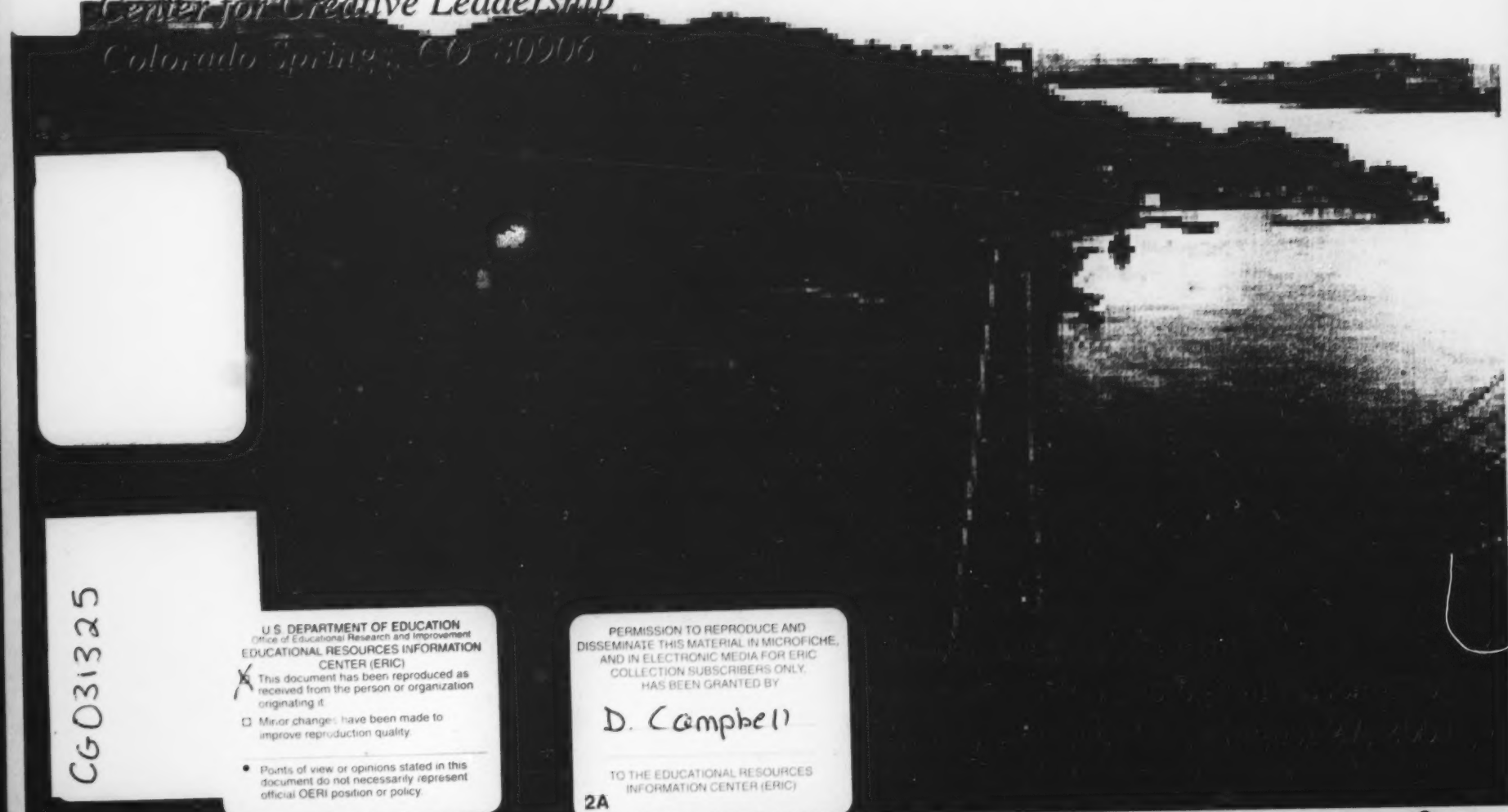
This paper presents a brief overview of the history of the Campbell Interest and Skill Survey, tracing its roots back to the early days of vocational interest inventories, beginning in the 1920s. Many of the important themes of the survey construction are discussed, such as the domains assessed; content of items; response format of items; scale construction; conceptual arrangements of the scoring scales; and interpretation methods. A description and evaluation is included of the 1992 addition to the Interest Inventory. (Contains 9 tables and 13 references.) (Author/JDM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 458 464

The History and Development of The Campbell Interest and Skill Survey (CISS)

David Campbell, Ph.D.
Center for Creative Leadership
Colorado Springs, CO 80906



CG031325

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY
D. Campbell
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
2A

BEST COPY AVAILABLE

The History and Development of The Campbell Interest and Skill Survey (CISS)

David Campbell, PhD
Center for Creative Leadership
Colorado Springs, CO 80906

Paper presented at the American
Psychological Association Convention,
San Francisco, California
August 27, 2001

Abstract

This is a brief overview of the history of the Campbell Interest and Skill Survey, tracing its roots back to the early days of vocational interests inventories, beginning in the 1920's. Many of the important themes of survey construction are discussed, such as the domains assessed, content of items, response format of items, scale construction, conceptual arrangements of the scoring scales, and interpretation methods. The 1992 addition of Skill items to the Interest Inventory is described and evaluated.

The Campbell Interest and Skill Survey (CISS) is a recent addition to a long line of statistically based career inventories, dating directly back to E. K. Strong, Jr who began this kind of assessment with his publication of the Strong Vocational Interest Blank (SVIB) in 1927. Strong's brilliant, break-through approach of actually collecting data from occupational samples and comparing each occupation's responses with those of a population reference sample to develop scoring scales has underlaid psychological assessment ever since. (Campbell, Hyne, and Nilsen, 1992), (Strong, 1927)

Six years later, in 1933, for reasons that he never wrote down anywhere, Strong created a second version specifically for women, at which point the two inventories became known as the SVIB-M and the SVIB-W. Strong was presumably reacting to the both the culture of the 1920's and '30's, which offered women somewhat different career paths than men, and also to the statistical reality that women responded differently to many of the items on the booklet, suggesting that different norms were necessary. Even so, there was a lot of overlap in his approach; sixty percent of the items on the two booklets were identical.

Strong slightly revised his Men's form in 1938, the Women's in 1946, which was also the year that he retired as Professor of Psychology from the Stanford University Business School. In 1955, in conjunction with a symposium held in his honor at the University of Minnesota, Strong began laying plans for a further revision of his two inventories, working in collaboration with Professor Ralph Berdie, Director of the large University of Minnesota Student Counseling Bureau (and Strong's son-in-law), and Professor Kenneth Clark, Chairman of the University's Psychology Department. Strong's immense and meticulously

organized occupational normative data files were transferred to Minnesota and prepared for computer input.

I became involved in 1959 when, as a graduate research assistant, I was hired to do the necessary computer analyses, and continued on in 1960 when I finished my PhD and joined the Minnesota faculty as an Assistant Professor of Psychology. In 1962, the University established the Center for Interest Measurement Research (CIMR), largely in anticipation of receiving the remainder of Strong's archival files, and I was appointed as CIMR's first Director.

In 1963, with perhaps one-half of the revision projects completed, Strong passed away, Clark had left Minnesota to become Dean of the College of Arts and Sciences at the University of Colorado, and Berdie was continuing with his substantial administrative duties in the University's Office of the Dean of Students; thus, the major responsibility for completing the revisions fell into my lap.

The revision of the Men's Form was published in 1966 (Campbell, 1966); the Women's form in 1969. The major changes in the revisions have been detailed in the Handbook for the SVIB, (Campbell, 1971), and a more informal history of the SVIB to this point appeared in 1967 (Campbell, 1967).

Although the revisions were technically well received, politically and sociologically they immediately became controversial. They were seen as vivid and visual evidence of the alleged chauvinistic nature of career inventories of the time, a viewpoint that was encouraged by an undoubtedly overly cute color coding scheme: the Men's form was printed on blue paper, the Women's on pink paper, an approach that was to lead to one of the most disastrous public relations fiascos in the history of psychological testing.

Over the next three years, the controversy escalated. We were briskly questioned by the President's Commission on the Equal Status of Women; the Wayne State University Faculty Senate passed a resolution condemning the SVIB; the American Personnel and Guidance Association (the predecessor of the American Counseling Association) sent a Task Force to Minneapolis to investigate the situation; the U. S. Office of Education sponsored a three-day conference on bias in career inventories; university students in classes and counseling sessions were continually raising questions; women professional colleagues were increasingly critical; close female friends and family members were not reticent in voicing their dismay; a Doonesbury cartoon commented sardonically on the segregated use of pink and blue forms during the law school orientation week for Joanie Caucus and her classmates; and lawsuits of various kinds were being threatened.

It became apparent that a major shift in strategy was necessary. In 1972, plans were made, then implemented, to combine the two forms into a single "unisex" booklet, utilizing the best features of each. This led to the publication of the Strong-Campbell Interest Inventory (SCII) in 1974 (Campbell, 1974). The tempest immediately receded; the changeover of forms was ninety percent complete in the first year of publication, an unheard speed for the adoption of a revision. Further modifications appeared over the next

decade, smoothing the changes and documenting their impact (Campbell, 1977; Campbell and Hansen, 1981; Hansen and Campbell, 1985).

Shortly after the revision was finished, in 1973-74, I left the University of Minnesota for a one year sabbatical leave at the newly founded Center for Creative Leadership in Greensboro, North Carolina, a non-profit, educational institution focused on the overlapping areas of creativity and leadership. During this year, I was recruited away from Minnesota and joined the Center as a Vice President, responsible for the Center's incipient research and training programs. From its inception, the Center has been a heavy user of psychological assessments, and I found myself on the other side of the desk, so to speak, as a user instead of a creator of surveys.

My interest in career assessment was still high, especially as I could now see directly the value of psychological assessment in leadership development, and in this applied setting, while still assessing the impact of the substantial changes that had been made in creating the Strong-Campbell, I began pondering what the next generation of career surveys should look like.

Because innumerable clients, classroom participants, and seminar attendees through the years had responded to their Strong-Campbell results with something like, "Well, I may be interested in that area (science, art, mechanical activities, sales, or whatever), but I don't have any skills there," it seemed promising to somehow quantify the respondents' self-reported skills along with their interests, both to ascertain if their self-assessment was accurate when using a more detailed, systematic approach, and to give them more specific guidance on what they might do next in a developmental sense.

Consequently, I began tinkering with a variety of methods to measure self-reported skills. Working at the Center was a huge advantage because I had daily contact with leaders, managers, executives, commanders, administrators – whatever term you wish to use – from a wide variety of organizations, and could easily try out new approaches.

My first attempt in about 1975, was to create a primitive, experimental listing termed "An Erratic List of Twenty-Five Skills." Some examples of the items were:

Give an interesting, informative after-dinner speech,
Factor quadratic equations.
Tape a bruised ankle.
Tell stories to children, keeping them enthralled.
Jump rope for three minutes without missing.
Conduct religious exercises.
Sketch portraits of people.
Appraise the value of jewelry.
Drill soldiers in close-order drill.

The initial item format called for a simple "Yes, I can do it. – No, I can't." response.

In the initial tallies, even with the primitive item format and casual "eyeball" data analysis techniques, there were sizeable and sensible differences between different occupational groups in what they said they could do well – sales personnel said they could sell, military commanders said they could drill troops, scientists said they could conduct experiments, artists said they could sketch pictures, police officers said they could handle dangerous situations, and so forth.

About this time, a family experience further convinced me that we were on the right track. My high school aged son completed the SAT, and received the report of his scores (which were good enough to get him into MIT). On his report form I noticed a simple grid reporting his self-evaluation on several global skills, such as "Scientific, Social, Athletic, Leadership, etc" (All of this is from memory, the form has been lost to history.) I was impressed by two things: first, his ruthless honesty, and second, the apparent validity of the pattern of his ratings; he rated himself pretty much as I would have.

"Interesting," I thought, "he knows that these results could influence his chances of admission, yet he has evaluated himself quite honestly, both in his strengths and his weaknesses."

Both of these experiences convinced me that the potential for a more formalized, standardized inventory of self-reported skills was a viable possibility. However, by then I was enmeshed in a demanding administrative role – Executive Vice-President of CCL – and so for the next few years, I merely continued the tinkering, imbued as I was with the Minnesota induced need to fool around with statistics.

Then an even more impactful, external event intruded to further slow progress, specifically an intense author-publisher clash regarding the Strong-Campbell Inventory. In 1983, with no warning and no prior consultation, I received a letter from the publisher, Stanford University Press, announcing that from then on a commercial publisher would take over the management of the inventory, including the research and development functions. When I protested, citing my contractual rights as author, Stanford Press brushed aside my concerns but reminded me of the so-called "non-compete" clause in our contract that forbid me from working on any competing work.

When negotiations proved futile, three years of protracted legal wrangling followed, culminating in 1987 with a decision from the Ninth Circuit Court of Appeals in San Francisco, reaffirming my rights as author but leaving the non-compete issue ambiguous. By this time, both sides were worn down by the legal posturing, and both sides realized that the author-publisher relationship had become so poisonous that future collaboration was impossible; consequently, we severed our relationship. Stanford Press took back all rights to the inventory, and I took back all rights to my name.

At the time, I considered it a huge personal and professional setback – I was being removed from any influence on an inventory that I had been focused on for two decades -- but, as with so many apparent disasters in life, it has proven to be a substantial blessing. This episode has allowed me to build a new inventory, starting with a clean slate without

the troublesome inertia of the past, and with the benefits of 25 years of experience with both survey creation and survey applications.

Problems Faced / Lessons Learned.

The construction of a new survey requires the solving of a series of interlocking problems, which solutions then become a series of "lessons learned" about the activity. This sequence of problems and subsequent learnings has been used to structure the remainder of this paper, although in actuality the sequence was considerably messier than it appears in this after-the-fact analysis.

What domain shall be used for conceptual guidance?

The content domain used for guidance in writing the skill items was the domain of vocational interests that has been carved out over the years by the cumulative findings of earlier researchers such as E. K. Strong, Jr., L.L. Thurstone, John Holland, and the earlier work on the Strong-Campbell. Basically, the domain is defined by the five to seven major dimensions of interests that have turned up repeatedly in factor studies of interests over the years. By 2001, this domain has become so familiar that it is difficult to remember how many earlier blind alleys were entered. For example, in his earliest attempts, Strong used items such as, ("How interested are you in working with...") "People with protruding jaws," "Cripples," "Bolshevists," and "Women who are cleverer than you are." (See copies of Strong's early forms in Campbell, 1971.)

He also included the item, "Negroes" in his first versions (1927 for Men, 1933 for Women) but removed it from the 1938 version for Men. Astonishingly, that item remained on the Women's version until it was removed in the 1969 revision. (Campbell, 1971)

Over time, this domain has become almost totally occupationally defined; there are no items such as, "How interested are you in understanding your emotions?," or "Would you rather spend free time alone or with a group?" For purposes of helping people make educational and career decisions, the vocabulary of the occupational world has proven to be powerful, and the factor structure for vocational interests provides the most efficient and focused guideline for studying occupational differences.

How shall the organizing factors be named?

John Holland's RIASEC terminology for his famous Six "Hexagonal" dimensions has become the standard of the field: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional. His conceptual brilliance in creating the circular hexagonal organization of the factors instead of Strong's and other researchers usual linear "top-to-bottom" listing is one of the true theoretical breakthroughs in the field.

However, our experience in working with clients has suggested that the Holland titles are sometimes confusing. For example, many lay people interpret the label "Investigative" as police work rather than as Holland's scientific activities. Further, even more troubling,

sometimes the labels are rejected, especially the "Conventional" label which is often seen as socially undesirable, especially in professional settings, covering as it does mainly clerical and secretarial activities. To avoid this problem in the new inventory, this dimension has been expanded to include mostly managerial and financial services and renamed "Organizing."

Additionally, our statistical solutions, along with intuition, suggested that Holland's Realistic factor should be divided in two dimensions: Producing -- mechanical, agricultural and construction activities, and Adventuring, -- police, military, and athletic activities.

Finally, using action verbs as labels for the dimensions seemed preferable to using Holland's adjectives. To that end, the labels for the CISS dimensions are (with the roughly comparable Holland dimensions in parentheses):

Influencing (Enterprising)
Organizing (Conventional)
Helping (Social)
Creating (Artistic)
Analyzing (Investigative)
Producing (Realistic)
Adventuring (Realistic).

Thus, the RIASEC acronym becomes IOHCNPA, with N used for Analyzing because the A is reserved for Adventuring. One possible mnemonic for this arrangement is "I Often Hear Counselors Naturally Prefer Anchovies.

How shall the items be worded?

For the CISS, two important changes in item phraseology were instituted as a result of experience with the earlier inventories. First, instead of listing just an occupational title, all occupations have been defined -- for example, "An Engineer, designing large building projects," and "A Physical therapist, helping people recover from injuries." This has served two purposes: it assures that all respondents are reacting to the same definition for the occupation, and it helps those respondents who are not quite certain of the nature of the activities of some of the lesser known occupations.

Second, a more complicated issue. Over the years in working with interest item statistics, it has become clear that the most powerful items in terms of separating occupational samples are occupational titles, as contrasted with school subjects, specific activities, or other categories. For example, the items "(Be an) artist" or "(Be a) research scientist" spread members of different occupations apart in their responses better than do items such as "Sketch pictures of people or places" or "Operate scientific equipment." The latter items are useful in expanding the bandwidth of item coverage and adding redundancy to scoring scales, but they are not as effective in discriminating between occupational samples. Consequently, a large number of occupational titles have been included among the interest items.

However, a different dynamic appeared when preparing the skill items. In early tryouts, respondents strongly resisted being asked to evaluate their skill level for specific occupations. While they were comfortable in evaluating their skill level for specific activities, such as sketching pictures or operating equipment, they rebelled against evaluating themselves more globally as artists or scientists. They complained, "There are just too many factors to be considered in making that judgment." Consequently, the skills item pool has no occupational titles.

Examples of CISS Skill Items appear in Table 1.

Table 1

Acting the lead role in a demanding drama.
Advising people making career decisions, helping them plan their future.
Analyzing data, using statistical concepts.
Being responsible for animals in a kennel or stable.
Cooking a gourmet meal, including several different courses.
Helping people plan their investment strategies.
Initiating new creative educational methods.
Managing a large forest preserve.
Overcoming the sales resistance of potential customers.
Presiding as master of ceremonies at a large program.

What item response format should be used?

Historically, many different item response formats have been used in various surveys. They can be grouped into two main categories: free response where the respondent is free to choose any one of the response options, and forced-choice where there is some constraint on the individual's response. The free response format includes the familiar "True-False" format used in the MMPI, the L-I-D (Like-Indifferent-Dislike) format that was used in the SVIB, SCII, and is now used in the SII, and a wide variety of "Likert" scales which tend to use five or seven point formats with the options labeled some version of "Strongly Agree" to "Strongly Disagree."

The forced choice format includes the format: "Here are two items: A and B. Mark which one you prefer." It is termed forced choice because the respondent does not have the option of liking both, or rejecting both. The MBTI has several items of this type, as did the earlier forms of the SVIB. Another forced choice format is the triad format where the individual is shown three items and asked to select the one most preferred and the one least preferred. This format is used in the Kuder Preference Record and in the Edwards Personal Preference Schedule.

Surely one of these formats must superior to the others but there are few, if any, studies with direct comparisons of various item formats. The reason is obvious and practical. No author is going to go through the enormous amount of work and expense

involved in developing a survey with one format, establishing all of the necessary descriptive validity and reliability statistics, and then repeating it all again with a different format, just to see which one is better.

Consequently, in the absence of direct comparative studies, what criteria should be used to decide which format is preferable to the others?

There are two major criteria: the first is the quality of statistical data produced. Some items formats are more stable than others, and some item formats produce more precise data than others. The second criteria is acceptance by respondents, which encompasses issues such as ease and accuracy of understanding the items, lack of irritation with the requested responses, lack of response bias (such as always picking the response on the left), and lack of extraneous idiosyncratic response patterns (such as always avoiding the extreme positions).

There is no simple way to know just how each of the possible item response formats meets each of these criteria. However, simultaneously with the development of the CISS, several other surveys were under parallel development, specifically surveys dealing with organizational climate (Campbell, 1990, 1994), leadership characteristics as seen by self and others (Campbell, 1991), the performance of teams (Campbell and Hallam, 1994), the characteristics of team leaders (Campbell and Hallam, 1997), and later, a community survey, which is as yet still unpublished. A firm strategy of establishing standardized characteristics across all of these surveys was followed to shorten training time for each survey; thus, each survey was to have the same item format, the same scale clustering techniques, the same norming methods, the same procedural checks for accuracy of processing, and the same profile layouts.

For experimental purposes, each survey went through several drafts during the tryout stages, and each draft was used with at least a few samples of respondents to see how it worked. Some of these drafts had different response formats; thus, there was a great deal of information available to guide the choice of the standardized characteristics.

In choosing a common item response format for all of these surveys, the decision-making process was especially thorough and cautious. Because throughout the entire published lifetime of the survey, each respondent will be confronted with the selected item response format and thus will provide data through the medium of this format, the choice of the final item format should be considered to be among the most important decisions. (Note that each respondent to the Strong is still presented with the L-I-D format first used by Strong in 1927. The dilemma is clear; changing this format would invalidate any future comparisons with previously collected data. In a vivid sense, the item response format holds hostage seventy years of comparative data.)

During the initial tryouts of these varied surveys, several item formats were experimented with, and a six point scale, which differs from most traditional response formats in that it has no mid-point, seemed to be the most useful in terms of creating stable, valid statistical data, and in being acceptable to respondents. Consequently, the final

choice for all of the surveys was a six point scale, ranging variously from "Strongly Like" to "Strongly Dislike," or "Always" to "Never," or "Strongly Agree" to "Strongly Disagree," as appropriate for the different item contents.

However, none of these choices worked grammatically with the Skill items. Further, we quickly learned that the biggest problem was to create an adequately high ceiling for the Skill items. Earlier projects had shown that respondents tend to report relatively positive views of their own skills such that, for example, approximately twenty-five percent of the population report that they are in the top ten percent on any given skill.

After more experimentation, we came up with the following six labels for the skill item responses; the numbers in parentheses after each one is the approximate percentage of the general employed population selecting each option across the entire Skills survey:

Expert – Widely recognized as excellent in this area.(10)
Good – Have well developed skills here. (20)
Slightly above average -- Average, or a touch above. (25)
Slightly below average -- Average, or a touch below. (20)
Poor -- Not very skilled here (15)
None -- Have no skills in this area. (10)

These response labels tend to produce a relatively normal, that is, a bell-shaped distribution of responses; thus, they were chosen for the final format.

One of the most desirable qualities of response formats is that they spread the responses over a wide range. "Thin split" item responses, that is, those where the responses are heavily bunched up on one response are undesirable because they cut down on the discriminating power of the inventory. Table 2 below demonstrates that the interest items do this quite well and, in a satisfying outcome, Table 3 demonstrates that the six labels used for the skill items also achieves this end. For the Skill items especially, a considerable number of attempts was necessary to achieve this result.

Table 2

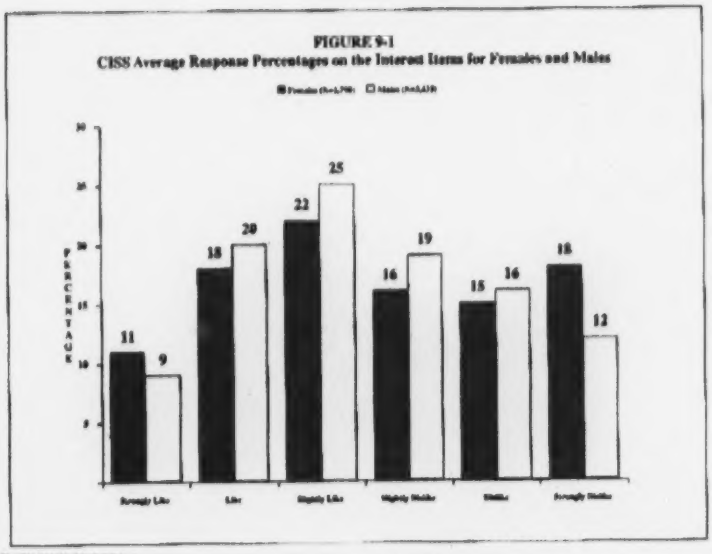
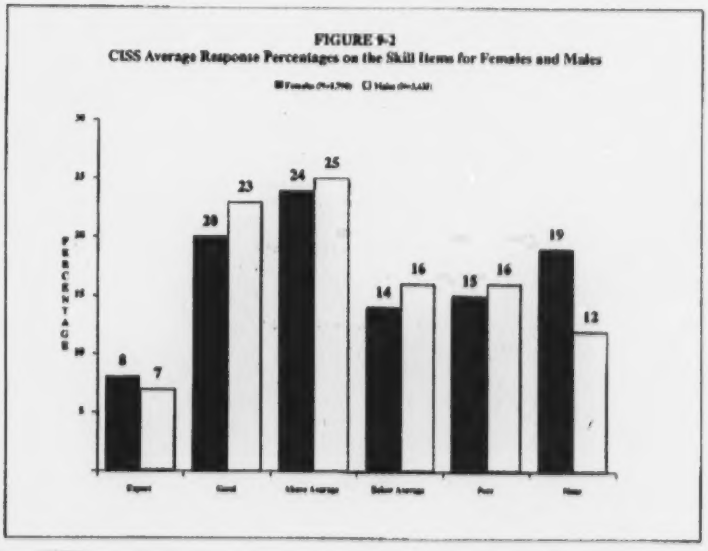


Table 3



How long should the survey be?

This has to be a judgment call as there is no practical way of testing out inventories of various lengths. Strong's first forms contained 420 and 410 items respectively for the Men's and Women's Forms, but he quickly shortened both of them to 400 items for the prosaic reason that that was the maximum number of items that could be handled by an IBM 805 Test Scoring Machine, which was the test scoring work horse of the 1930's. The

BEST COPY AVAILABLE

1960's revision continued in that vein with the Men's Form having 399 items and the Women's 398. The numbers were changed slightly, simply to make it easier to distinguish between the answer sheets.

Working with inventories of this length for several years taught us three things: first, there is a fair amount of respondent resistance to long inventories, second, improved item selection techniques can help weed out weaker items with little loss to the inventory overall, and third, an improved item response format produces more robust scales even with fewer items per scale. Consequently, in 1971 the Strong-Campbell was reduced to 325, and in 1992 the CISS interest item pool was reduced even further to 200 items, allowing the addition of 120 skill items. Thus, the CISS has a total of 320 items, covering both interests and skills.

Are the items free from any real or perceived bias?

In developing any new contemporary inventory, exquisite attention must be paid to the presence of any real or perceived bias in the items, such as with racial or gender themes. If even one item offends any well-defined demographic group, the inventory will come under attack. From the extensive, earlier experience with the Strong inventories, we learned to avoid not only sexist and racist items but also to be sensitive to a wide range of other concerns, such as socio-economic class and geographic issues. For example, the item, "Belong to a country club" was dropped as being elitist, and the label on the front of the inventory booklet asking for "Zip Code" has been changed to "Postal Code," in response to international sensitivities.

An anecdote to demonstrate how difficult it is to avoid all linguistic pitfalls: A last minute change in the CISS occurred when an alert copyreader called attention to a potential problem. At this final stage in 1991, the inventory had been through several iterations, it had been widely used to collect occupational normative data from several thousand respondents, at least a dozen experts had looked at it for one reason or another, and yet this potential problem had slipped through. The item in question asked the respondent to indicate their skill level for the item, "Charming members of the opposite sex."

Data had already indicated that this item was more often selected by people in occupations where a strong sense of gregariousness is useful, such as realtor or interior decorator.

Fortunately, almost ten years ago when the CISS was first published, when such sensitivities were just emerging, the antennae of the copyreader detected the possible controversial problem of sexual orientation, and the item was changed to "Charming and entertaining other people." There has not been a single complaint. Indeed, there have been no complaints of bias against any of the CISS items.

How shall the items be grouped into scoring scales?

There are four types of scoring scales on the CISS:

- Orientation Scales, which are homogeneous scales focused on the major factor dimensions,
- Basic Scales, which are essentially sub-scales of the Orientation scales and more tightly focused on specific areas,
- Occupational Scales which reflect the dominant interests and skills of specific occupations, and
- Special Scales, Academic Focus and Extroversion Introversion.

Working from considerable experience with earlier surveys, the Orientation and Basic Scales were created by clustering together items that had relatively high statistical intercorrelations; thus, they are termed "homogeneous scales." They are easy to understand and interpret because their content is obvious; each scale covers a well-defined area such as science or art or sales.

The Occupational Scales were created by comparing specific samples with a general reference group, such as comparing architects with People-in-general, and aggregating the items that the architects answered quite differently than People-in-general into the Architect Scale. The Special Scales were constructed in a similar empirical fashion, comparing people who had achieved different levels of academic achievement for the Academic Focus Scale, and comparing externally defined extraverts and introverts for the Extroversion Scale. Thus, these scales are termed "empirical" scales.

The homogeneous scales can be quite short, even as few as five or six items, because they are so tightly focused. The occupational scales are much more heterogeneous; thus they need to be longer to achieve the same level of stability as the short homogeneous scales, at least 25 or 30 items in length, and longer is even better if sufficient discriminating items are available.

One dramatically different dynamic turned up with the skill items in contrast with the interest items. From the very beginning, starting with Strong's work in the 1920's, both Like and Dislike responses have been scored on the interest scales, and indeed it has always been apparent that respondents will strongly avoid occupational settings where they must perform specific activities that they find distasteful; thus, Dislike responses are important. Consequently if a respondent shares the same likes AND dislikes as, say, architects, he or she will score high on the Architect scale and be more inclined to enter, and remain in that occupation.

As we were making a deliberate attempt to keep the construction of the skill scales identical to the construction of the interest scales, we followed the same scale construction methods. However, midway in the process, we noticed that we had one occupational sample that had no positive weights for any of the skill items, which meant that,

statistically, there was not a single skill item that the people in that occupation felt that they performed better on than People-in-General, a surprising finding.

This led us to think through the conceptual approach of using negatively weighted skill items on any of the occupational scales. While it seems quite logical to counsel someone that they might consider architecture as a career because they share the dislikes as well as the likes of architects, it somehow seems less appropriate to counsel people to choose a career direction because they share the incompetencies of the people in that career. It would be especially unsettling to say to someone, "You have indicated that you have no notable skills; therefore you should consider this occupation where you at least share their incompetencies."

Consequently, a policy decision was made to exclude all negatively weighted skill items on all of the occupational scales. (Later, a second occupational sample also turned up with the same pattern, that is, no single skill item on which they as a group thought they were better than average.)

How shall the scoring scales be normed?

The norming process has been well established by decades of earlier work on vocational interest inventories. First, some sensible general sample is selected to serve as a reference group. This sample is then scored with the scoring scales, and the resulting statistics are used to set the population norms. For the CISS, the reference group raw scores were converted to standard scores so that this sample has a mean of 50 and a standard deviation of 10 on all of the scales. These scores are usually called T-scores.

Every subsequent respondent is scored with this same conversion so that the meaning of a score is immediate apparent. For example, a score of 60 is always one standard deviation about the population mean.

For the CISS, there was early concern about the effect of age on these norms. The plan was for the norms to be established on adults, but then to be used for comparisons with students. As it seems intuitively clear that adults have more skills than students, it initially appeared that the norms would be too high, and perhaps a second set of norms only for students would have to be established.

However, after the normative data had been collected, and after some students had been compared with the norms, it became clear that this was not a problem; students tended to respond to the individual items much as did adults. Vivid data to support this conclusion appears in Table 4 where the percentages of various age groups selecting each of the six possible responses to each of the items on the inventory are graphed. Clearly there is no need for separate norms.

In studying just why this is the case, it became clear that respondents read the items in a manner relevant to their particular situation. For example, in response to the item, "Can act the lead role in a demanding drama," high school students tend to think of class plays

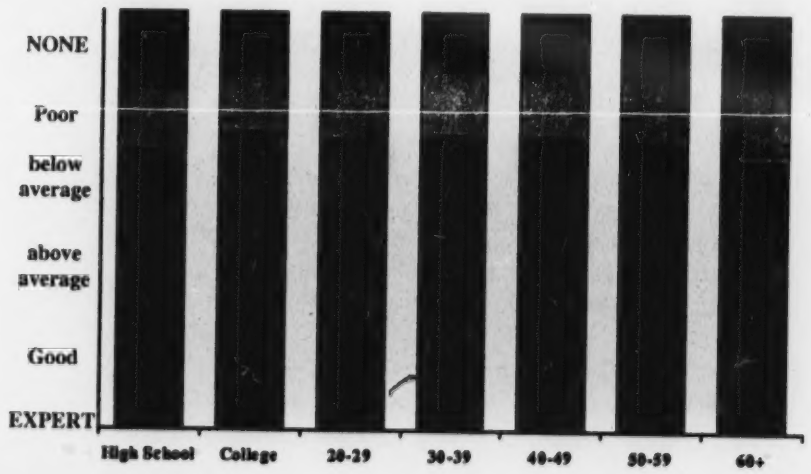
while adults tend to think of either community or professional theater, depending on their own experience.

Similarly, in response to the item, "Can inspire my teammates to superior performance," high school athletes tend to think of teams at their level, college athletes tend to think of teams at the collegiate level, and adults tend to think of whatever level of team is relevant for them.

Consequently, no separate age norms were necessary.

Table 4

Response Percentage Frequencies for Skill Items by Age



How to document validity?

The main question of validity for the skill scales is whether or not respondents in different occupations report different patterns of skills, and the answer to that is unequivocal – they clearly do. Furthermore, the nature of the differences is quite sensible, as can be seen in the following examples.

The Manual for the CISS reports validity data for all of the scales; only three examples will be given here: 1) for the Counseling Basic Scale, 2) for the Attorney Occupational Scale, and 3) for the Academic Focus Scale. In the relevant tables that follows, data for both Interest and Skill Scales are plotted graphically which makes it possible to visually compare the relative power of the Interests and Skill to separate occupational samples.

For illustrative purposes, the items on the Counseling Basic Scales are listed below.

Table 5

Illustrative items on the Counseling Basic Scales

Positively Weighted Interest Items

A career counselor, helping people make important career decisions.
A psychologist, studying human behavior.
Study Group Dynamics
Soothe angry or upset people.

Positively Weighted Skill Items

Advising people making career decisions...
Aiding an angry husband and wife to understand each other better.
Constructing a psychological analysis of someone...
Drawing people out in conversations...

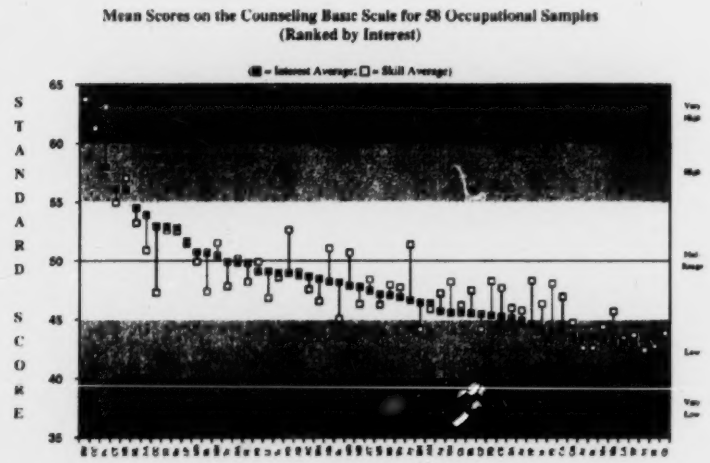
Mean scores on these Counseling Interest and Skill Scales are plotted in Table 6 for fifty-eight occupational samples by rank order of each sample's Interest Scale mean. The Interest mean is represented by a solid square, the Skill mean by a hollow square. The five highest and five lowest occupations are listed in the Table. The other samples are identified by initials at the bottom of the Table; these initials are identified in the CISS Manual.

For the Counseling Scales, the five occupations scoring highest were quite predictable: Social Workers, Guidance Counselors, Psychologists, Corporate Trainers and Human Resource Directors – all occupations involved in one way or another with counseling and human development activities. Many of them undoubtedly have degrees in counseling psychology.

The five occupational samples scoring low have less obvious cohesiveness among them, but they are clearly not "people occupations."

Table 6

Mean Scores on the Counseling Basic Scales for 58 Occupational Samples



Analogous data are reported below for the Attorney Occupational Scales. For illustrative purposes, some representative items from the Attorney's Scales are shown in Table 7. Each of these items substantially discriminated between attorneys, both male and female, and the general reference sample.

Table 7

Illustrative Items on the Attorney Occupational Scale

Positively Weighted Interest Items

A diplomat, negotiating agreements between countries.
An elected official, developing new public programs.
A judge, presiding over a courtroom
A trial lawyer, arguing cases in court.

Negatively Weighted Interest Items

An engineer, designing large building projects.
A factory superintendent...
An interior designer, planning room layouts.
A technical writer, preparing scientific manuals.

Positively Weighted Skill Items

Acting the lead role in a demanding drama.
Debating issues in a public forum.

BEST COPY AVAILABLE

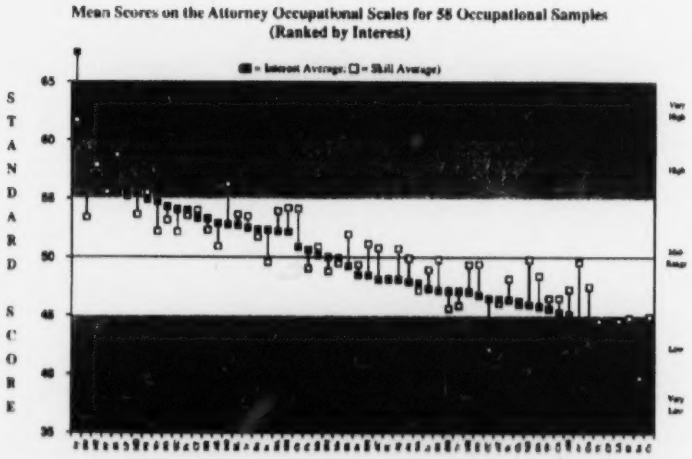
Negotiating compromises between conflicting parties.
Preparing detailed financial contracts...

No negatively weighted skill items were used on any of the scales, as discussed above.

Mean scores on the Attorney scales are presented in Table 8 for 58 occupational samples, and the five highest and lowest scoring occupational samples are listed in the Table. For this scale, the five highest scoring occupations were not quite so obvious and, with the exception of the high-scoring attorneys, their scores were considerably lower than the mean scores on the Counseling Scales above. The explanation is almost certainly that we have surveyed only one "legal oriented" occupational sample. If we had surveyed other samples such as Trial Lawyers, District Attorneys, or Courtroom Judges, they almost certainly would also have scored high.

Table 8

Mean Scores on the Attorney Interest and Skill Scale
for 58 Occupational Samples



As can be seen, both the Interest and Skill scales work well in terms of spreading occupations apart; the occupations scoring high and the occupations scoring low are sensible, and the differences between them are large enough to be practically meaningful. (The scales are normed so that the general population means and standard deviations are 50 and 10 for each scale, respectively.)

However, the Skill scales do not separate occupations quite as well as do the Interest Scales. The evidence for this conclusion is that there was a strong tendency for the

BEST COPY AVAILABLE

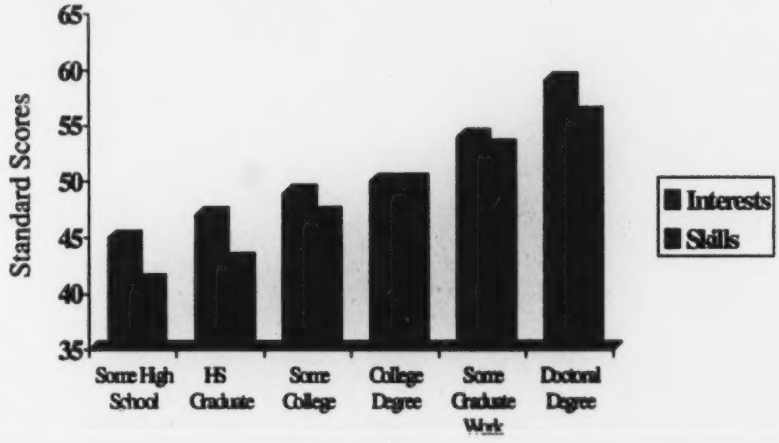
occupational means of the Skill scales to fall between the general reference mean (50) and the occupational mean of the Interest scales. What this means in a practical sense is that one can make a slightly better prediction of someone's future career direction by asking them, "What do you like to do?" versus asking them, "What are you capable of doing?" though both answers are clearly relevant.

Table 9 contains a slightly different kind of validity data, specifically the mean scores of respondents on the Academic Focus scales who have earned different levels of formal degrees. Again, the sequence of means was quite tidy, ranging over 15 points, one and one-half standard deviation (42 to 57) from those respondents with no high school diploma to those respondents with doctoral level degrees. (The specific sample N's for each of these individual groups are not available, but the aggregated samples total in the thousands.)

Once again, the combination of item content and scale construction methods led to measures with satisfactory validity.

Table 9

Average Academic Focus Scores by Level of Education



How can the "on diagonal" pairs of interest and skill scores be interpreted versus the "off diagonal" pairs?

In using the CISS in pre-publication try-out counseling sessions, one of the intriguing findings was that the combination of the two scores for each scale, that is, the Interest score and the Skill score, was more important clinically than either scale taken alone. Consequently, a two – by – two labeling approach was developed, with each pair of scores falling into either a definite "on diagonal" pattern, that is, with both scores either high or

low, or "off diagonal" pattern, that is, with one score high, the other low, or visa versa. The four quadrants were labeled and interpreted as follows:

Interests high, Skills high = **Pursue**: These are areas of both high interest and high confidence. Include in top career choices.

Interests high, Skills low = **Develop:** These are areas of high interest, low confidence. Either accept as a vocational, or pursue skill building.

Interest low, Skills high = **Explore:** These are areas of low interest but high confidence. Reconsider why interest is low (there may be some unnecessary biases operating) or transfer skills to another area.

Interests low, Skills low = **Avoid:** These are areas of both aversion and low confidence. Exclude from top career choices.

Patterns where both scores fell in the middle were left unlabeled, and interpreted merely as "Mid-range" with no special significance.

Curiously, this labeling approach, which has nothing to do with the usual statistical concerns of survey construction, has proven to be quite useful when working with counselees where psychological interpretations are more important than purely analytical considerations.

<u>How shall the scales be interpreted in terms of Actual Skills versus Self-perceived Skills?</u>

Easily the most surprising and most useful finding to come out of this lengthy focus on measured, self-reported skills is that it may be more useful conceptually to think of these scales as measures of self-confidence than of actual skills. While self-confidence in the activity, and the actual skill appear to be closely related, where deviations occur, "self-confidence," or lack thereof, seems to be the dominant force.

There are no specific data available to back up these assertions, only a string of episodes, case studies, anecdotes, and persuasive examples. From the frenetic MBA student who marked 78 percent "Expert," to the beaten down nurse, going through a divorce, on probation in her job, distressed over her son's drug problem, who marked 75 percent "None", case after case suggest that one's perceived level of skills is influenced by one's confidence is applying these skills. This is not to say that it is all perception; the sensible rank-ordering of occupational mean scores strongly suggest that actual skills are being evaluated by respondents. Still, especially at the extremes, perception becomes reality.

From a counselor's standpoint, this is probably the most important learning. An individual's mental state appears to have a more direct impact on their perceived competence, and thus on the activities that they are likely to engage in, than does a more analytical assessment. One of the counselor's main tasks may well be to help the individual reach a realistic appraisal of their competencies; this is especially important for

those individuals whose self-criticism is holding them back from moving in directions where they can succeed beyond their expectations.

This is not to encourage false self-esteem. The abilities truly have to be there, but the measured approach suggested here can help an individual deal more directly with inaccurate personal assessment.

Clinical Insights

Three of the most important clinical insights that have come from twenty-five years of experience with the assessment of self-reported skills are:

First, skills can be measured in the same manner that interests have been measured for some seventy years, and the resulting measurements are related to occupational choices. Thus, much of what has been learned about the measurement of interests can be generalized to the measurement of skills.

Second, these skill measurements are sensible and useful as early as the high school years.

Third, the standardized measures of self-reported skills appear to track some intertwined combination of the actual skill itself and the respondent's sense of self-confidence in carrying out the skill. At the negative extreme of the distributions of responses, experience suggests that about five percent of the reference population, which is essentially employed adults, have substantially more skills that they report on the CISS and, conversely, at the other end of the distribution, again based on experience, another five percent appear almost manic in reporting their skills as extremely wide-spread. Both of these extremes appear to be more related to either an under- or over-abundance of self-confidence than to the presence of the actual skill.

These "clinical insights" can perhaps provide fruitful suggestions for the next wave of inquiry into the measurement of skills.

References

Campbell, D.P. (1966) Manual for the Strong Vocational Interest Blank, Stanford University Press, Stanford, CA.

Campbell, D.P. (1967) The Strong Vocational Interest Blank: 1927 – 1967, Chapter appearing in McReynolds, P., (1967) Advances in Psychological Assessment, Science and Behavior Books, Palo Alto, CA

Campbell, D.P. (1971) Handbook for the Strong Vocational Interest Blank, Stanford University Press, Stanford, CA.

Campbell, D.P. (1974) Manual for the Strong-Campbell Interest Inventory, Stanford University Press, Stanford, CA

Campbell, D.P. (1977) Manual for the Strong-Campbell Interest Inventory (2nd Edition), Stanford University Press, Stanford, CA

Campbell, D.P. and Hansen, J.C. (1981) Manual for the Strong-Campbell Interest Inventory (3rd Edition), Stanford University Press, Stanford, CA

Campbell, D.P. (1990) Manual for the Campbell Organizational Survey, National Computer Systems, Minneapolis, MN

Campbell, D.P. (1991) Manual for the Campbell Leadership Index, National Computer Systems, Minneapolis, MN.

Campbell, D.P, Hyne, S.A., and Nilsen, D.L. (1992) Manual for the Campbell Interest and Skill Survey, National Computer Systems, Minneapolis, MN.

Campbell, D.P. and Hallam, G., (1994) Manual for the Team Development Survey, (1994), National Computer Systems, Minneapolis, MN

Hallam, G and Campbell, D.P. (1999) User's Guide for the Team Leader Profile, National Computer Systems, Minneapolis, MN

Hansen, J.C. and Campbell, D.P. (1985) Manual for the Strong Interest Inventory (4th Edition), Stanford University Press, Stanford, CA

Strong, E.K., Jr. (1927) A vocational interest test. The Educational Record, 8, 107-121

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of these documents from another source, please provide the following information regarding the availability of these documents. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Counseling & Student Services
University of North Carolina at Greensboro
201 Ferguson Building
PO Box 26171
Greensboro, NC 27402-6171